JOURNAL OF MULTIVARIATE ANALYSIS 13, 234–256 (1983)

Measuring the Efficiency of Trigonometric Series Estimates of a Density

PETER HALL

Australian National University

Communicated by E. J. Hannan

Some types of density estimators, particularly those based on trigonometric series, converge reasonably quickly to their limit except in the neighbourhood of one or two singularities. In this situation the mean integrated square error, the traditional measure of the efficiency of a density estimator, is an unsatisfactory measure. The notion of partial mean integrated square error is introduced and used to compare the performance of trigonometric series estimators. The results lead to consideration of some new estimators which have excellent properties from the points of view of both efficiency and ease of computation.

1. INTRODUCTION

Let $X_1, X_2, ..., X_n$ be a sample of independent random variables from the distribution with density f. If the support of f is confined to $[-\pi, \pi]$, it is possible to estimate f by a weighted trigonometric series,

$$\hat{f}_n(x; m) = (2\pi)^{-1} \left[1 + 2 \sum_1^m w_m(j)(\hat{a}_j \cos jx + \hat{b}_j \sin jx)\right],$$

where $\hat{a}_j = n^{-1} \sum_{i=1}^n \cos jX_i$, $\hat{b}_j = n^{-1} \sum_{i=1}^n \sin jX_i$, and the $w_m(j)$ are constant weights. The study of such orthogonal series estimators dates from the early papers of Cencov [5], van Ryzin [12], and Schwartz [13]. Kronmal and Tarter [9] considered in some detail the special case where each $w_m(j) = 1$, and also discussed the choice of more general weight functions. Watson [20] showed that in a certain sense the optimal form of the weights depends on the unknown density itself, and so can only seldom be determined. He suggested that in practice the simplest choice, $w_m(j) = 1$

Received August 20, 1980; revised October 21, 1980.

AMS 1970 subject classification: Primary 62G05.

Key words: density estimate, mean square error, mean integrated square error, singular integral, trigonometric series.

0047-259X/83 $3.00

for all j, might be close to optimal. More recent discussions of trigonometric series estimators, including weighted series estimators, are contained in Anderson and de Figueiredo [1], Tarter [14], Tarter and Raman [15], Wahba [16, 17], and Walter and Blum [19].

It is common when studying a trigonometric series estimator to measure its "efficiency" by its mean integrated square error (MISE),

$$\int_{-\pi}^{\pi} E[\hat{f}_n(x) - f(x)]^2 \, dx.$$

Indeed, the MISE criterion was used by Watson [20] to determine the optimal form for the weights. The bias of $\hat{f}_n$, however, can be considerably greater towards the endpoints of $[-\pi, \pi]$ than anywhere else, and so the MISE can provide an unduly pessimistic measure of the performance of $\hat{f}_n$ over much of the interval of estimation. In the present paper we propose an alternative to the MISE, which we term the *partial* mean integrated square error (PMISE). The PMISE measures the performance of an estimator over "most" of the interval $[-\pi, \pi]$, and is often of a smaller order of magnitude than the MISE.

In Section 2 we shall demonstrate the utility of the PMISE criterion by applying it to some simple, well-known trigonometric series estimators. In Section 3 we introduce weighted trigonometric series estimators based on periodic singular integrals. The PMISEs of these estimators converge at the same rate as for a kernel estimator, and more quickly than for many of the well-known trigonometric series estimators. Indeed, if the weights are chosen correctly, then the PMISE of a periodic singular integral estimator will be 5.5% less than that of the best kernel estimator. The weights are defined in an absolute sense, and do not depend on the unknown density. Tarter and Raman [15] have shown that it is possible to construct weighted trigonometric series estimators which improve on kernel estimators, although their choice of weights depended on the density to be estimated, and they used the MISE to compare performances.

Let us introduce the concept of PMISE by considering the simple Fourier series estimator $\hat{f}_n(x; m)$ defined above, in the special case where each $w_m(j) = 1$. For a suitable choice of $m = m(n)$ the rate of convergence of the mean square error (MSE), $E[\hat{f}_n(x) - f(x)]^2$, is of $O(n^{-2/3})$ at each point $x \in (-\pi, \pi)$, but this rate is not attained uniformly on $(-\pi, \pi)$; see [8]. It is not difficult to see, however, that the MSE does converge uniformly on $[-\pi + \varepsilon, \pi - \varepsilon]$ for each $\varepsilon > 0$. This disparity is caused by "edge effects," or Gibbs' phenomenon, in the bias of $\hat{f}_n(x; m)$; see Carslaw [4, Chap. IX] or Lanczos [10, pp. 217–219]. The bias of the estimator will not converge to zero towards the ends of the interval unless $f(\pi-) = f(-\pi+)$. Indeed, if the values of $f(\pi-)$ and $f(-\pi+)$ do not coincide, then the MISE may not converge at a faster rate than $O(n^{-1/2})$ [8]. Therefore the MISE does not satisfactorily measure the performance of the estimator on almost all of $[-\pi, \pi]$. (The "edge effects" would apply with equal validity to any point of discontinuity of f within $[-\pi, \pi]$.)

It is common for statisticians to make the assumption that a density assumes common values at the ends of its range (Elderton [6, p. 41]), although such an assumption is not always convenient. For example, if we wish to use a trigonometric series to estimate a density over only a subset (a, b) of its range, the values of the density at the endpoints a and b would very likely be unknown. Kronmal and Tarter's [9] examples on the use of trigonometric series estimators are of just this type; see [9, Sect. 6]. The fact that an estimator performs badly at the ends of an interval, however, need not be a drawback. To obtain an accurate estimate of f at the endpoints of (a, b) we would extend our interval to $(a - \varepsilon, b + \varepsilon)$, and condition on observations from our sample lying in this set. This technique would allow us to obtain an estimate of f whose MSE converged at a rate of $O(n^{-2/3})$ uniformly on (a, b). In conducting this procedure it is unnecessary to make assumptions about the values of f at the endpoints of intervals.

Considerations of this type lead us to the following definition of PMISE: given two numbers ε_1 and ε_2 in the range $0 < \varepsilon_i < \pi$, set

$$J(\varepsilon_1, \varepsilon_2) = \int_{-\pi+\varepsilon_1}^{\pi-\varepsilon_2} E[\hat{f}_n(x) - f(x)]^2 \, dx.$$

In the case of many estimators the pathological behaviour at the endpoint π is the same as that at $-\pi$, and so there is no real loss of generality in setting $\varepsilon_1 = \varepsilon_2 = \varepsilon$. We shall always make this assumption, and so define the PMISE by $J(\varepsilon, \varepsilon)$. Walter and Blum [19] and Walter [18] have presented an ingeneous general approach to the theory of generalized kernel estimates, but they assume that the density and its derivatives are continuous at the endpoints of the interval of estimation. The concept of PMISE is needed to adequately describe the contrary case, and just this case arises when a density is estimated over a proper subset of its support.

2. Classical Trigonometric Series Estimates

Let $X_1, X_2, \ldots, X_n$ be independent observations of a distribution with a density f of bounded variation, and define

$$\hat{a}_j = \hat{c}_j = n^{-1} \sum_{i=1}^{n} \cos jX_i \qquad \text{and} \qquad \hat{b}_j = \hat{s}_j = n^{-1} \sum_{i=1}^{n} \sin jX_i.$$

If f has its support confined to $(-\pi, \pi)$ and is continuous there, then

$$\hat{f}_n(x; m) = (2\pi)^{-1}\left[1 + 2\sum_{1}^{m}(\hat{a}_j \cos jx + \hat{b}_j \sin jx)\right]$$

is an estimate of $f(x)$. If the support is restricted to $(0, \pi)$, two estimates of $f(x)$ are

$$\hat{f}_{n1}(x; m) = \pi^{-1}\left(1 + 2\sum_{1}^{m}\hat{c}_j \cos jx\right) \qquad \text{and} \qquad \hat{f}_{n2}(x; m) = 2\pi^{-1}\sum_{1}^{m}\hat{s}_j \sin jx.$$

THEOREM 1. *Suppose f has its support confined to $(-\pi, \pi)$, and $f(-\pi) \equiv f(-\pi+)$ and $f(\pi) \equiv f(\pi-)$ are well defined. If f' is of bounded variation on $(-\pi, \pi)$, then for any $0 < \varepsilon < \pi$,*

$$\int_{-\pi+\varepsilon}^{\pi-\varepsilon} E[\hat{f}_n(x; m) - f(x)]^2\, dx$$

$$= (m/n\pi)\int_{-\pi+\varepsilon}^{\pi-\varepsilon} f(x)\, dx + (2\pi^2 m^2)^{-1}[f(\pi) - f(-\pi)]^2$$

$$\times \int_{\varepsilon}^{\pi} (1 - \cos x)^{-1}\, dx + O(n^{-1}\log m) + O(m^{-5/2})$$

as m and $n \to \infty$. If f has a bounded second derivative on $(-\pi, \pi)$, then the term $O(n^{-1}\log m)$ may be replaced by $O(n^{-1})$.

THEOREM 2. *Suppose f has its support confined to $(0, \pi)$. If f' is of bounded variation on $(0, \pi)$, and if $f(0) \equiv f(0+)$ and $f(\pi) \equiv f(\pi-)$ are well defined, then for any $0 < \varepsilon < \pi/2$,*

$$\int_{\varepsilon}^{\pi-\varepsilon} E[\hat{f}_{n2}(x; m) - f(x)]^2\, dx$$

$$= (m/n\pi)\int_{\varepsilon}^{\pi-\varepsilon} f(x)\, dx + (\pi^2 m^2)^{-1}[f(0)^2 + f(\pi)^2]$$

$$\times \int_{\varepsilon}^{\pi-\varepsilon} (1 - \cos x)^{-1}\, dx + O(n^{-1}\log m) + O(m^{-5/2})$$

as m and $n \to \infty$. If f'' is bounded on $(0, \pi)$, then the term $O(n^{-1}\log m)$ may be replaced by $O(n^{-1})$. If f'' is of bounded variation on $(0, \pi)$, and if $f'(0) \equiv f'(0+)$ and $f'(\pi) \equiv f'(\pi-)$ are well defined, then for any $0 < \varepsilon < \pi/2$,

$$\int_{\varepsilon}^{\pi-\varepsilon} E[\hat{f}_{n1}(x;m)-f(x)]^2\,dx = (m/n\pi)\int_{\varepsilon}^{\pi-\varepsilon} f(x)\,dx$$
$$+(\pi^2 m^4)^{-1}[f'(0)^2+f'(\pi)^2]$$
$$\times\int_{\varepsilon}^{\pi-\varepsilon}(1-\cos x)^{-1}\,dx+O(n^{-1}+m^{-9/2})$$

as m and $n\to\infty$.

It follows that if we choose $m=O(n^{1/3})$, then the PMISEs of $\hat{f}_n(x;m)$ and $\hat{f}_{n2}(x;m)$ converge at a rate of $O(n^{-2/3})$, and if we choose $m=O(n^{1/5})$, then the PMISE of $\hat{f}_{n1}(x;m)$ converges at a rate of $O(n^{-4/5})$, which is the same rate as for a kernel estimator. The *minimal* PMISE of the cosine series estimator, however, has the form

$$\min_m \int_{\varepsilon}^{\pi-\varepsilon} E[\hat{f}_{1n}(x;m)-f(x)]^2\,dx$$
$$\sim 5\pi^{-6/5}2^{-8/5}\left[\int_{\varepsilon}^{\pi-\varepsilon} f(x)\,dx\right]^{4/5}$$
$$\times\left\{[f'(0)^2+f'(\pi)^2]\int_{\varepsilon}^{\pi-\varepsilon}(1-\cos x)^{-1}\,dx\right\}^{1/5} n^{-4/5},$$

as $n\to\infty$, provided that the right-hand side does not vanish. This quantity increases like $1/\varepsilon^{1/5}$ as $\varepsilon\to 0$, and so a larger increase in the sample size is needed to improve the performance of the estimator towards the endpoints of $(-\pi,\pi)$ than is needed to improve its performance at the centre of the interval. A kernel estimator does not usually suffer from this deficiency. The singular integral estimators we shall introduce in Section 3 perform better than the cosine series estimator in this respect.

Proof of Theorem 1. Let a_j, b_j and a_j', b_j' be the Fourier cosine, sine coefficients of f and f', respectively (e.g., $a_j=\int_{-\pi}^{\pi}\cos ju\, f(u)\,du$). Then

$$a_j=-j^{-1}b_j', \qquad b_j=(-1)^{j+1}j^{-1}[f(\pi)-f(-\pi)]+j^{-1}a_j',$$

and if $|x|<\pi$,

$$\pi b(x;m)=\pi E[f(x)-\hat{f}_n(x;m)]$$
$$=\sum_{m+1}^{\infty}(a_j\cos jx+b_j\sin jx)$$
$$=[f(\pi)-f(-\pi)]\sum_{m+1}^{\infty}(-1)^{j+1}j^{-1}\sin jx$$
$$+\sum_{m+1}^{\infty}j^{-1}(a_j'\sin jx-b_j'\cos jx).$$

Since f' is bounded and satisfies Dirichlet's conditions, then (see [4, p. 269])

$$\int_{-\pi+\varepsilon}^{\pi-\varepsilon}\left[\sum_{m+1}^{\infty} j^{-1}(a_j' \sin jx - b_j' \cos jx)\right]^2 dx \leqslant \int_{-\pi}^{\pi}$$

$$= \sum_{m+1}^{\infty} j^{-2}(a_j'^2 + b_j'^2) = O(m^{-3}).$$

Therefore by the triangle inequality for an L^2 norm,

$$A_m = \left\{\pi^2 \int_{-\pi+\varepsilon}^{\pi-\varepsilon} b(x;m)^2\, dx\right\}^{1/2}$$

$$= \left\{[f(\pi) - f(-\pi)]^2 \int_{\varepsilon}^{2\pi-\varepsilon}\left[\sum_{m+1}^{\infty} j^{-1}\sin jx\right]^2 dx\right\}^{1/2} + O(m^{-3/2}).$$

Using Abel's method of summation we find that

$$\sum_{m}^{\infty} j^{-1}\sin jx = \sum_{m}^{\infty} j^{-1}(j+1)^{-1}\, S_j(x) - m^{-1}S_{m-1}(x), \tag{1}$$

where

$$S_n(x) = \sum_{1}^{n} \sin jx = \{\sin nx + \sin x - \sin[(n+1)x]\}/2(1-\cos x).$$

For a constant C depending only on ε,

$$\int_{\varepsilon}^{2\pi-\varepsilon}\left[\sum_{m}^{\infty} j^{-1}(j+1)^{-1}(\sin jx - \sin(j+1)x)\right]^2 (1-\cos x)^{-2}\, dx$$

$$\leqslant C\int_0^{2\pi}\left[\sum_{m}^{\infty} j^{-1}(j+1)^{-1}(\sin jx - \sin(j+1)x)\right]^2 dx$$

$$= C\pi\left[2\sum_{m}^{\infty} j^{-2}(j+1)^{-2} - \sum_{m}^{\infty} j^{-1}(j+1)^{-2}(j+2)^{-1}\right.$$

$$\left. - \sum_{m+1}^{\infty} j^{-2}(j^2-1)^{-1}\right] = O(m^{-3}). \tag{2}$$

Since $\sum_{m+1}^{\infty} j^{-1}(j+1)^{-1} = m^{-1} + O(m^{-2})$, then

$$A_m = \left\{[f(\pi) - f(-\pi)]^2 \int_{\varepsilon}^{2\pi-\varepsilon} [\sin x(2m(1-\cos x))^{-1}\right.$$

$$\left. - m^{-1}S_m(x)]^2\, dx\right\}^{1/2} + O(m^{-3/2})$$

$$= \left\{(8m^2)^{-1}[f(\pi)-f(-\pi)]^2 \int_{\varepsilon}^{2\pi-\varepsilon} (1-\cos x)^{-2}[2-2\cos x - \cos 2mx \right.$$

$$\left. - \cos(2m+2)x + 2\cos(2m+1)x]\, dx\right\}^{1/2} + O(m^{-3/2})$$

$$= \left\{(4m^2)^{-1}[f(\pi)-f(-\pi)]^2 \int_{\varepsilon}^{2\pi-\varepsilon} (1-\cos x)^{-1}\, dx\right\}^{1/2} + O(m^{-3/2})$$

$$= \left\{\tfrac{1}{2}m^{-2}[f(\pi)-f(-\pi)]^2 \int_{\varepsilon}^{\pi} (1-\cos x)^{-1}\, dx\right\}^{1/2} + O(m^{-3/2}). \tag{3}$$

Next we observe that

$$\hat{f}_n(x; m) = n^{-1} \sum_{i=1}^{n} D_m(x - X_i),$$

where $D_m(x) = \sin[(2m+1)x/2]/2\pi \sin(x/2)$ is the Dirichlet kernel, and so

$$n \operatorname{var}[\hat{f}_n(x; m)] = \int_{-\pi}^{\pi} D_m^2(x-u) f(u)\, du + O(1)$$

uniformly in $|x| \leqslant \pi - \varepsilon$. Let $g(x, u) = [f(u) - f(x)]/(u - x)$. Then

$$A = \sup_{|x| \leqslant \pi-\varepsilon,\, |u| \leqslant \pi} |g(x, u)| < \infty$$

and

$$\int_{-\pi}^{\pi} D_m^2(x-u) f(u)\, du = f(x) \int_{-\pi}^{\pi} D_m^2(u)\, du + r_m,$$

where

$$|r_m| \leqslant (2\pi/\varepsilon) A \int_{-\pi}^{\pi} |u|\, D_m^2(u)\, du.$$

Now,

$$D_m^2(u) = (2\pi)^{-2}(2m+1) \left\{1 + 2 \sum_{1}^{2m} [1 - j/(2m+1)] \cos ju\right\}$$

and so

$$(2m+1)^{-1}\,4\pi^2\int_{-\pi}^{\pi}|u|\,D_m^2(u)\,du=\pi^2-8\sum_0^{m-1}(2j+1)^{-2}$$
$$+\,8(2m+1)^{-1}\sum_0^{m-1}(2j+1)^{-1}.$$

Since $8\sum_0^\infty(2j+1)^{-2}=\pi^2$, then

$$\int_{-\pi}^{\pi}|u|\,D_m^2(u)\,du=O(\log m).$$

If f'' is bounded on $(-\pi,\pi)$, then with

$$h(x,u)=[f(u)-f(x)-(u-x)f'(x)]/(u-x)^2$$

we have

$$B=\sup_{|x|\leqslant\pi-\varepsilon,\,|u|\leqslant\pi}|h(x,u)|<\infty.$$

Therefore,

$$\int_{-\pi}^{\pi}D_m^2(x-u)f(u)\,du=f(x)\int_{-\pi}^{\pi}D_m^2(u)\,du+s_m,$$

where

$$|s_m|\leqslant(2\pi/\varepsilon)^2B\int_{-\pi}^{\pi}u^2D_m^2(u)\,du=O(1).$$

Since $\int_{-\pi}^{\pi}D_m^2(u)\,du=(2m+1)/2\pi$, then

$$\begin{aligned} n\operatorname{var}[\hat f_n(x;m)]-mf(x)/\pi &= O(\log m), &&\text{if } f' \text{ is bounded},\\ &= O(1), &&\text{if } f'' \text{ is bounded},\end{aligned}$$

uniformly in $|x|\leqslant\pi-\varepsilon$, completing the proof.

Proof of Theorem 2. We may write

$$c_j=\int_0^{\pi}\cos ju\,f(u)\,du=j^{-2}[(-1)^jf'(\pi)-f'(0)]-j^{-2}\int_0^{\pi}\cos ju\,f''(u)\,du$$

and

$$s_j=\int_0^{\pi}\sin ju\,f(u)\,du=-j^{-1}[(-1)^jf(\pi)-f(0)]+j^{-1}\int_0^{\pi}\cos ju\,f'(u)\,du.$$

Therefore

$$\begin{aligned}
\tfrac{1}{2}\pi b_i(x; m) &= \tfrac{1}{2}\pi[f(x) - E\hat{f}_{ni}(x; m)] \\
&= \sum_{m+1}^{\infty} j^{-2}[f'(\pi)\cos j(x+\pi) - f'(0)\cos jx] \\
&\quad - \sum_{m+1}^{\infty} j^{-2} c_j'' \cos jx, \qquad \text{if} \quad i = 1, \\
&= \sum_{m+1}^{\infty} j^{-1}[f(0)\sin jx - f(\pi)\sin j(x+\pi)] \\
&\quad + \sum_{m+1}^{\infty} j^{-1} c_j' \sin jx, \qquad \text{if} \quad i = 2. \qquad (4)
\end{aligned}$$

If $d_i(x; m)$ denotes the last series in these expressions, then under the conditions of the theorem,

$$\begin{aligned}
\int_0^{\pi} d_i(x; m)^2\,dx &= O(m^{-5}), \qquad \text{if} \quad i = 1, \\
&= O(m^{-3}), \qquad \text{if} \quad i = 2.
\end{aligned} \qquad (5)$$

Furthermore,

$$\sum_{m}^{\infty} j^{-2}\cos jx = \sum_{m}^{\infty} j^{-2}(j+1)^{-2}(2j+1)\,C_j(x) - m^{-2}C_{m-1}(x), \qquad (6)$$

where

$$C_n(x) = \sum_{0}^{n} \cos jx = \{\cos nx - \cos x - \cos[(n+1)x] + 1\}/2(1 - \cos x),$$

and also

$$\begin{aligned}
&\int_{\varepsilon}^{2\pi-\varepsilon} \left[\sum_{m+1}^{\infty} j^{-2}(j+1)^{-2}(2j+1)(\cos jx - \cos(j+1)x)\right]^2 \\
&\qquad \times (1-\cos x)^{-2}\,dx = O(m^{-5}).
\end{aligned}$$

Combining this with (4)–(6) we see that

$$\begin{aligned}
A_{m1} &= \left\{\tfrac{1}{4}\pi^2 \int_{\varepsilon}^{\pi-\varepsilon} b_1(x; m)^2\,dx\right\}^{1/2} \\
&= \left\{\tfrac{1}{4}m^{-4} \int_{\varepsilon}^{\pi-\varepsilon} e_1(x; m)^2\,dx\right\}^{1/2} + O(m^{-5/2}),
\end{aligned}$$

where

$$e_1(x; m) = f'(\pi)(-1)^m[\cos(m+1)x + \cos mx](1+\cos x)^{-1} + f'(0)[\cos(m+1)x - \cos mx](1-\cos x)^{-1},$$

and combining (1), (2), (4), and (5) we deduce that

$$A_{m2} = \left\{\tfrac{1}{4}\pi^2 \int_\varepsilon^{\pi-\varepsilon} b_2(x; m)^2\, dx\right\}^{1/2} = \left\{\tfrac{1}{4}m^{-2} \int_\varepsilon^{\pi-\varepsilon} e_2(x; m)^2\, dx\right\}^{1/2} + O(m^{-3/2}),$$

where

$$e_2(x; m) = f(\pi)(-1)^m[\sin(m+1)x + \sin mx](1+\cos x)^{-1} + f(0)[\sin(m+1)x - \sin mx](1-\cos x)^{-1}.$$

After some manipulation we find that

$$\int_\varepsilon^{\pi-\varepsilon} e_i(x; m)^2\, dx = e_i \int_\varepsilon^{\pi-\varepsilon} (1-\cos x)^{-1}\, dx + O(m^{-1}),$$

where $e_1 = f'(0)^2 + f'(\pi)^2$ and $e_2 = f(0)^2 + f(\pi)^2$. Therefore

$$A_{mi} = \left\{\tfrac{1}{4}m^{-r(i)} e_i \int_\varepsilon^{\pi-\varepsilon} (1-\cos x)^{-1}\, dx\right\}^{1/2} + O(m^{-r(i)/2-1/2}), \tag{7}$$

where $r(1) = 4$ and $r(2) = 2$. Next observe that

$$\hat{f}_{n1}(x; m) = n^{-1} \sum_{i=1}^{n} [D_m(x - X_i) + D_m(x + X_i)]$$

and

$$\hat{f}_{n2}(x; m) = n^{-1} \sum_{i=1}^{n} [D_m(x - X_i) - D_m(x + X_i)],$$

from which we see that

$$n \operatorname{var} \hat{f}_{ni}(x; m) = E[D_m(x - X_1) \pm D_m(x + X_1)]^2 + O(1) = \int_0^\pi D_m^2(x-u) f(u)\, du + O(1)$$

uniformly in $\varepsilon \leqslant x \leqslant \pi - \varepsilon$, since

$$\int_0^\pi D_m^2(x+u) f(u)\, du \leqslant \pi \sup_{0<u<\pi} D_m^2(x+u) = O(1)$$

uniformly in $\varepsilon \leqslant x \leqslant \pi - \varepsilon$. The proof may be completed as in the case of Theorem 1 if we observe that

$$\int_0^\pi D_m^2(x-u)\, du = \int_{-\pi}^\pi D_m^2(u)\, du + r_m,$$

where

$$|r_m| \leqslant \int_{\{\varepsilon < |u| \leqslant \pi\}} D_m^2(u)\, du = O(1).$$

To obtain our last results in this section we consider the Fejér form of the Fourier series estimate. The Fejér form of the cosine series estimate was considered in [8], and in this case it may be seen that the bias converges to zero uniformly on $(0, \pi)$ at a rate of $O(m^{-1})$. Therefore the MISE provides an accurate picture of the rate of convergence of this estimator on any subinterval of $(0, \pi)$. This is not the case, however, for the Fejér form of $\hat{f}_n(x; m)$, which we define by

$$\begin{aligned} \hat{f}_n^*(x; m) &= (m+1)^{-1} \sum_{j=0}^m \hat{f}_n(x; j) \\ &= (2\pi)^{-1} \left\{ 1 + 2 \sum_1^m [1 - j/(m+1)](\hat{a}_j \cos jx + \hat{b}_j \sin jx) \right\}. \end{aligned}$$

If $b_n^*(x; m) = f(x) - E[\hat{f}_n^*(x; m)]$ denotes the bias of this estimator, then

$$\int_{-\pi}^\pi b^*(x; m)^2\, dx = 2[f(\pi) - f(-\pi)]^2/m\pi + o(m^{-1})$$

(Hall [8]), although for any $\varepsilon > 0$,

$$\int_{-\pi+\varepsilon}^{\pi-\varepsilon} b^*(x; m)^2\, dx = m^{-2} \int_{-\pi+\varepsilon}^{\pi-\varepsilon} |(\tilde{f})'(x)|^2\, dx + o(m^{-2}),$$

as we shall shortly prove. Therefore the concept of PMISE is needed to adequately measure the performance of $\hat{f}_n^*$ on any interval contained inside $(-\pi, \pi)$.

The function $\tilde{f}$, the Hilbert transform of f, is the 2π-periodic function defined by the integral

$$\tilde{f}(x) = (2\pi)^{-1} \int_0^\pi [f(x-u) - f(x+u)] \cot(u/2)\, du,$$

where we extend f from $(-\pi, \pi]$ to $(-\infty, \infty)$ by periodicity. Note that the functions $(\tilde{f})'$ and $(f')^{\sim}$ are identical if and only if $f(\pi) = f(-\pi)$, and so we must take care to distinguish between them. Indeed, this point underlines the difference between the behaviour of the Fejér forms of the cosine and Fourier series expansions. The cosine series expansion of a density f with support confined to $(0, \pi)$ is the Fourier series expansion of the function $g(u) = f(u)$ if $u \geqslant 0$, $f(-u)$ if $u < 0$. Obviously $g(\pi) = g(-\pi)$. Therefore the cosine series expansion of f converges faster than the Fourier series expansion of

$$h(x) = 2f(2x - \pi),$$

unless $f(0) = f(\pi)$. (The function h is of course the density we would study if we were to use a direct Fourier series on $(-\pi, \pi)$ to estimate f.)

THEOREM 3. *Let f be a density with support confined to $(-\pi, \pi)$. Suppose f is bounded and continuous on $(-\pi, \pi)$, $\tilde{f}$ and $(\tilde{f})'$ exist and are bounded on $[-\pi + \varepsilon, \pi - \varepsilon]$ for each $\varepsilon > 0$, and*

$$\sup_{|x| \leqslant \pi - \varepsilon} \int_0^\pi u^{-2} |\tilde{f}(x+u) - \tilde{f}(x-u) - 2u(\tilde{f})'(x)|\, du < \infty$$

for each $\varepsilon > 0$. If $|x| < \pi$ and $f(x) \neq 0 \neq (\tilde{f})'(x)$, then

$$E[\hat{f}_n^*(x; m) - f(x)]^2 \sim mf(x)/3n\pi + m^{-2} |(\tilde{f})'(x)|^2$$

as m and $n \to \infty$. Furthermore, for any $0 < \varepsilon < \pi$,

$$\int_{-\pi+\varepsilon}^{\pi-\varepsilon} E[\hat{f}_n^*(x; m) - f(x)]^2\, dx \sim m \int_{-\pi+\varepsilon}^{\pi-\varepsilon} f(x)\, dx/3n\pi$$
$$+ m^{-2} \int_{-\pi+\varepsilon}^{\pi-\varepsilon} |(\tilde{f})'(x)|^2\, dx,$$

provided that the terms on the right do not vanish.

Some aspects of this theorem are similar to results which exist in the literature, but there does not seem to be one which exactly suits our present purpose. For example, we could use [3, Corollary 9.2.9, p. 346] to obtain an expression for the bias of $\hat{f}_n^*$, but it would be necessary to impose the condition that $f(-\pi+) = f(\pi-)$.

The conditions imposed on f in the theorem are satisfied if $(\tilde{f})'$ and $(\tilde{f})''$ are bounded on each interval $[-\pi+\varepsilon, \pi-\varepsilon]$, and $\int_{-\pi}^{\pi} |\tilde{f}(x)|\, dx < \infty$. The reasonableness of these conditions is best illustrated by an example. Suppose

$$f(x) = 3(x-\pi)^2/8\pi^3, \qquad -\pi < x < \pi.$$

Note that $f(-\pi+) \neq f(\pi-)$. The Fourier expansion is given by

$$8\pi^3 f(x)/3 = (4\pi^2/3) + 4\sum_{1}^{\infty} (-1)^j j^{-2} \cos jx + 4\pi \sum_{1}^{\infty} (-1)^j j^{-1} \sin jx,$$

and so

$$\begin{aligned} 8\pi^3 \tilde{f}(x)/3 &= 4\sum_{1}^{\infty} (-1)^j j^{-2} \sin jx - 4\pi \sum_{1}^{\infty} (-1)^j j^{-1} \cos jx \\ &= 4\sum_{1}^{\infty} (-1)^j j^{-2} \sin jx + 4\pi \log[2\cos(x/2)]. \end{aligned}$$

Hence

$$8\pi^3 (\tilde{f})'(x)/3 = -4\log[2\cos(x/2)] - 2\pi \tan(x/2)$$

and

$$8\pi^3 (\tilde{f})''(x)/3 = 2\tan(x/2) - \pi \sec^2(x/2).$$

Therefore the conditions on f and $\tilde{f}$ are all satisfied. Note that

$$\int_{-\pi}^{\pi} |(\tilde{f})'(x)|^2\, dx = \infty,$$

and so the results of Theorem 3 do not extend to the case $\varepsilon = 0$.

We conclude this section with a proof of our last result.

Proof of Theorem 3. We make use of the formula

$$f(x) - E[\hat{f}_n^*(x; m)] = \int_0^{\pi} [\tilde{f}(x+u) - \tilde{f}(x-u)]\, \kappa_m(u)\, du,$$

where

$$\kappa_m(u) = \sin[(m+1)u]/4\pi(m+1)\sin^2(u/2).$$

(See Zygmund [21, p. 91].) Therefore we may write

$$\Delta_m(x) = f(x) - E[\hat{f}_n^*(x; m)] - 4(\tilde{f})'(x) \int_0^\pi \sin(u/2)\, \kappa_m(u)\, du$$

$$= \int_0^\pi g(x, u) \sin[(m+1)u]\, du/4\pi(m+1),$$

where

$$g(x, u) = [\tilde{f}(x+u) - \tilde{f}(x-u) - 4(\tilde{f})'(x)\sin(u/2)]\sin^2(u/2).$$

The Riemann–Lebesgue lemma implies that if $|x| < \pi$, then $\Delta_m(x) = o(m^{-1})$, and since

$$\int_0^\pi \sin(u/2)\, \kappa_m(u)\, du \sim 1/4m$$

as $m \to \infty$, then

$$f(x) - E[\hat{f}_n^*(x; m)] \sim m^{-1}(\tilde{f})'(x) + o(m^{-1}) \tag{8}$$

for each fixed x. The bounded convergence theorem implies that

$$\int_{-\pi+\varepsilon}^{\pi-\varepsilon} \left| \int_0^\pi g(x, u) \sin[(m+1)u]\, du \right|^2 dx \to 0$$

as $m \to \infty$, and so

$$\int_{-\pi+\varepsilon}^{\pi-\varepsilon} [f(x) - E\hat{f}_n^*(x; m) - m^{-1}(\tilde{f})'(x)]^2\, dx = o(m^{-2}). \tag{9}$$

The results (8) and (9) account for the bias terms of the expansions in Theorem 3. To obtain the variance, observe that

$$\hat{f}_n^*(x; m) = n^{-1} \sum_1^n F_m(x - X_i),$$

where $F_m(x) = \{\sin[(m+1)x/2]/\sin(x/2)\}^2/2\pi(m+1)$ is the Fejér kernel. Therefore

$$n \operatorname{var}[\hat{f}_n^*(x; m)] = \int_{-\pi}^{\pi} F_m^2(x-u) f(u)\, du + O(1) = f(x) \int_{-\pi}^{\pi} F_m^2(u)\, du + o(m)$$

$$= f(x)\, m/3\pi + o(m)$$

uniformly in $|x| \leqslant \pi - \varepsilon$. Theorem 3 follows immediately.

3. Singular Integral Estimates

Suppose density f has its support confined to $(-\pi, \pi)$. Let $\{K_r\}$ be a sequence of 2π-periodic kernels satisfying conditions (10)–(12). For each r,

$$\int_{-\pi}^{\pi} K_r(u)\, du = 1, \tag{10}$$

$$\lim_{r \to r_0} \int_{-\pi}^{\pi} |K_r(u)|\, du = 1. \tag{11}$$

For all $\delta > 0$,

$$\lim_{r \to r_0} \int_{\{\delta < |u| < \pi\}} |K_r(u)|\, du = 0. \tag{12}$$

Define $\hat{f}_n(x) = n^{-1} \sum_{i=1}^{n} K_r(x - X_i)$. If we extend f from $(-\pi, \pi]$ to $(-\infty, \infty)$ by periodicity, then we have

$$E[\hat{f}_n(x)] = \int_{-\pi}^{\pi} f(x-u)\, K_r(u)\, du,$$

which is immediately recognized as a singular integral of f. It is to be expected that if $r = r(n) \to r_0$ at a suitable rate, then $\hat{f}_n(x; r)$ will be a consistent estimator of $f(x)$. Our first result makes this notion precise.

Theorem 4. *Let $x \in (-\pi, \pi)$ be a continuity point of the bounded density f, suppose $f(x) > 0$ and $r = r(n) \to r_0$ as $n \to \infty$. Then $\hat{f}_n(x; r) \to^p f(x)$ if and only if for all $\varepsilon > 0$,*

$$\int_{\{|K_r(u)| > n\varepsilon\}} |K_r(u)|\, du \to 0 \qquad \textit{and} \qquad n^{-1} \int_{\{|K_r(u)| \leqslant n\varepsilon\}} K_r(u)^2\, du \to 0,$$

where the integrals are taken over a single period.

In order to obtain mean square consistency we ask that

$$k_r = \int_{-\pi}^{\pi} K_r(u)^2\, du < \infty \qquad \text{for all} \quad r, \tag{13}$$

and

$$\int_{\{\delta < |u| < \pi\}} K_r(u)^2\, du = o(k_r) \qquad \text{as} \quad r \to r_0 \quad \text{for all} \quad \delta > 0. \tag{14}$$

It is obvious that both k_r and $\sup_u |K_r(u)|$ tend to infinity as $r \to r_0$. For many kernels they increase at the same rate, and we might ask that there be a constant C such that

$$\sup_u |K_r(u)| \leqslant Ck_r \qquad \text{for all} \quad r. \tag{15}$$

(The Fejér kernel F_m, defined during the proof of Theorem 3, satisfies conditions (10)–(12) and (13)–(15). The Dirichlet kernel D_m, however, does not satisfy condition (11).)

THEOREM 5. *If* (13) *and* (14) *hold, then under the conditions of Theorem* 1, $\hat{f}_n(x; r) \to f(x)$ *in mean square if and only if* $n^{-1}k_r \to 0$. *If* (13)–(15) *hold, the following three conditions are equivalent*:

$$n^{-1}k_r \to 0; \tag{16}$$

$$E[\hat{f}_n(x; r) - f(x)]^2 \to 0; \qquad \textit{and} \qquad \hat{f}_n(x; r) \xrightarrow{p} f(x). \tag{17}$$

If in addition

$$\sum_{n=1}^{\infty} \exp(-n\varepsilon/k_r) < \infty \qquad \textit{for all} \quad \varepsilon > 0, \tag{18}$$

then $\hat{f}_n(x; r) \to^{\text{a.s.}} f(x)$. *Finally, if* (13)–(15) *hold and* $n^{-1}k_r \to 0$, *then* $(n/k_r)^{1/2}[\hat{f}_n(x; r) - E\hat{f}_n(x; r)]$ *is asymptotically normal* $N(0, 1)$.

Proof of Theorem 4. Let X be a variable with density f and set $\mu_r(x) = E[\hat{f}_n(x; r)]$. Since

$$P(|K_r(x - X) - \mu_r(x)| > n\varepsilon) \leqslant (n\varepsilon)^{-1} \left[(\sup f(u)) \int_{-\pi}^{\pi} |K_r(u)| \, du + |\mu_r(x)| \right]$$
$$\to 0, \tag{19}$$

then it follows from Petrov [11, Theorem 3, p. 260] that $\hat{f}_n(x; r) \to^p f(x)$ if and only if for all $\varepsilon > 0$,

$$nP(|K_r(x - X) - \mu_r(x)| > n\varepsilon) \to 0, \tag{20}$$

$$E[\{K_r(x - X) - \mu_r(x)\} \, I(|K_r(x - X) - \mu_r(x)| \leqslant n\varepsilon)] \to 0 \tag{21}$$

and

$$n^{-1}E[\{K_r(x - X) - \mu_r(x)\}^2 \, I(|K_r(x - X) - \mu_r(x)| \leqslant n\varepsilon)] \to 0. \tag{22}$$

Let $E_n = \{|K_r(x - X) - \mu_r(x)| > n\varepsilon\}$. Conditions (10) and (11) imply that $E[\{K_r(x - X)\}^-] \to 0$, and (19) and (21) that $E[K_r(x - X) \, I(E_n)] \to 0$.

Therefore $E[|K_r(x-X)|\, I(E_n)] \to 0$, and so (20)–(22) are equivalent to the following two conditions: For all $\varepsilon > 0$,

$$\Delta_{n1} = E[|K_r(x-X)|\, I(|K_r(x-X)| > n\varepsilon)] \to 0,$$

and

$$\Delta_{n2} = n^{-1} E[|K_r(x-X)|^2\, I(|K_r(x-X)| \leqslant n\varepsilon)] \to 0.$$

For any $\delta > 0$,

$$\Delta_{n1} = \left[\int_{\{|K_r(u)| > n\varepsilon;\ |u| \leqslant \delta\}} + \int_{\{|K_r(u)| > n\varepsilon;\ |u| > \delta\}}\right] |K_r(u)| f(x-u)\, du$$

and

$$\Delta_{n2} = n^{-1} \left[\int_{\{|K_r(u)| \leqslant n\varepsilon;\ |u| \leqslant \delta\}} + \int_{\{|K_r(u)| \leqslant n\varepsilon;\ |u| > \delta\}}\right] |K_r(u)|^2 f(x-u)\, du,$$

where the integrals are over a single period and f is extended to $(-\infty, \infty)$ by periodicity. The second term in each of these expressions is dominated by a constant multiple of

$$[\sup_u f(u)] \int_{\{|u| > \delta\}} |K_r(u)|\, du,$$

which converges to zero. Choose δ so that for constants c_1 and c_2, $0 < c_1 < f(x-u) < c_2$ whenever $|u| \leqslant \delta$. Then Δ_{n1} and $\Delta_{n2} \to 0$ if and only if

$$\int_{\{|K_r(u)| > n\varepsilon;\ |u| \leqslant \delta\}} |K_r(u)|\, du \to 0$$

and

$$n^{-1} \int_{\{|K_r(u)| \leqslant n\varepsilon;\ |u| \leqslant \delta\}} |K_r(u)|^2\, du \to 0,$$

which conditions are equivalent to those of Theorem 4.

Proof of Theorem 5. If (13) and (14) hold, then $\operatorname{var}[\hat{f}_n(x; r)] \sim n^{-1} k_r f(x)$, proving the first part of the theorem. The second part will follow if we show that the last part of (17) implies (16). Now,

$$k_r \leqslant \int_{\{|K_r(u)| \leqslant n\}} K_r(u)^2\, du + [\sup_u |K_r(u)|] \int_{\{|K_r(u)| > n\}} |K_r(u)|\, du,$$

so that if (13)–(15) hold we may deduce from Theorem 4 that $n^{-1}k_r \leqslant o(1) + o(n^{-1}k_r)$, proving (16). If (18) holds, then strong consistency may be proved using the Borel–Cantelli lemma and the techniques leading to Bosq and Bluez's [2] Lemma 1. The last result of Theorem 5 follows directly from Lindeberg's theorem.

THEOREM 6. *Suppose each K_r is even and nonnegative, and that in addition to* (10)–(12) *the kernels satisfy* (13), (14), *and*

$$\lim_{r \to r_0} \left[\int_0^{\pi} (1 - \cos 2u)\, K_r(u)\, du\right] \Big/ \left[\int_0^{\pi} (1 - \cos u)\, K_r(u)\, du\right] = 4. \quad (23)$$

Let f be a density with support confined to $(-\pi, \pi)$ and with two bounded continuous derivatives on $(-\pi, \pi)$. If $-\pi < x < \pi$ and $f(x) \neq 0 \neq f''(x)$, then

$$E[\hat{f}_n(x; r) - f(x)]^2 \sim n^{-1}k_r f(x) + \left[f''(x) \int_{-\pi}^{\pi} (1 - \cos u)\, K_r(u)\, du\right]^2$$

as $n \to \infty$ and $r \to r_0$. Furthermore, for any $0 < \varepsilon < \pi$,

$$\begin{aligned} &\int_{-\pi+\varepsilon}^{\pi-\varepsilon} E[\hat{f}_n(x; r) - f(x)]^2\, dx \\ &\quad \sim n^{-1}k_r \int_{-\pi+\varepsilon}^{\pi-\varepsilon} f(x)\, dx \\ &\qquad + \left[\int_{-\pi+\varepsilon}^{\pi-\varepsilon} f''(x)^2\, dx\right]\left[\int_{-\pi}^{\pi} (1 - \cos u)\, K_r(u)\, du\right]^2, \end{aligned} \quad (24)$$

provided that the terms on the right do not vanish.

Condition (23) is satisfies by many kernel sequences, although not of course by the Fejér sequence. If K_r admits the expansion

$$K_r(u) = (2\pi)^{-1}\left[1 + \sum_1^{\infty} w_r(j) \cos ju\right]$$

for real numbers $w_r(j)$, then (23) holds if and only if $[1 - w_r(2)]/[1 - w_r(1)] \to 4$ as $r \to r_0$.

The expansions for the MSE and PMISE are reminiscent of their analogues in the case of a kernel estimator, since the bias is roughly proportional to $f''(x)$. Like the case of the kernel estimator, the optimal rate of convergence of the MSE and PMISE is $O(n^{-4/5})$. Note that expansion (24) is not necessarily valid in the case $\varepsilon = 0$.

Proof of Theorem 6. The techniques of the proof are well known. We define

$$\Delta_r(x) = E[\hat{f}_n(x; r)] - f(x) - f''(x) \int_{-\pi}^{\pi} (1 - \cos u) K_r(u)\, du$$

$$= \int_0^{\pi} [f(x+u) + f(x-u) - 2f(x) - 2(1 - \cos u) f''(x)] K_r(u)\, du.$$

Given ε and $\delta > 0$ we may choose $0 < \eta < \varepsilon$ such that whenever $|x| \leqslant \pi - \varepsilon$ and $|u| \leqslant \eta$,

$$|f(x+u) + f(x-u) - 2f(x) - 2(1 - \cos u) f''(x)| \leqslant \delta(1 - \cos u).$$

Therefore if $|x| \leqslant \pi - \varepsilon$,

$$|\Delta_r(x)| \leqslant \delta \int_0^{\eta} (1 - \cos u) K_r(u)\, du + B \int_{\eta}^{\pi} K_r(u)\, du,$$

where $B = 2 \sup_{|x| < \pi} [f(x) + |f''(x)|]$. Now,

$$\int_{\eta}^{\pi} K_r(u)\, du \leqslant (1 - \cos \eta)^{-2} \int_{\eta}^{\pi} (1 - \cos u)^2 K_r(u)\, du$$

$$\leqslant \tfrac{1}{2}(1 - \cos \eta)^{-2} \int_0^{\pi} [4(1 - \cos u) - (1 - \cos 2u)] K_r(u)\, du.$$

It follows from (23) that if r is sufficiently large,

$$|\Delta_r(x)| \leqslant 2\delta \int_0^{\pi} (1 - \cos u) K_r(u)\, du.$$

This inequality holds uniformly in $|x| \leqslant \pi - \varepsilon$, and takes care of the bias term in the expansions of Theorem 6. It is easily shown that

$$n \operatorname{var}[\hat{f}_n(x; r)] = k_r f(x) + o(k_r)$$

uniformly in $|x| \leqslant \pi - \varepsilon$, completing the proof.

Examples of Kernels Satisfying (10)–(15) *and* (23)

Let I denote the range of values of r. We refer the reader to [3] for the theory of singular integrals.

(i) *Rogosinski singular integral.* Let $I = \mathbb{Z}^+$ and $r_0 = \infty$, and define

$$K_m(u) = (2\pi)^{-1} \left\{ 1 + 2 \sum_1^m \cos[j\pi/(2m+1)] \cos ju \right\} \geqslant 0.$$

For this estimator $k_m \sim m/2\pi$ and $\int_{-\pi}^{\pi} (1 - \cos u) K_m(u)\, du \sim \pi^2/8m^2$.

(ii) *Korovkin singular integral.* Let $I = \mathbb{Z}^+$ and $r_0 = \infty$, and define

$$\begin{aligned} K_m(u) &= (2\pi)^{-1} \left[1 + 2 \sum_1^m w_m(j) \cos ju \right] \\ &= \frac{2 \sin^2[\pi/(m+2)]}{2\pi(m+2)} \left(\frac{\cos[(m+2)u/2]}{\cos[\pi/(m+2)] - \cos u} \right)^2, \end{aligned}$$

where

$$\begin{aligned} 2(m+2) \sin[\pi/(m+2)]\, w_m(j) = (m - j + 3) &\sin[(j+1)\pi/(m+2)] \\ - (m - j + 1) &\sin[(j-1)\pi/(m+2)]. \end{aligned}$$

Here $k_m \sim m(2 + 15/\pi^2)/12\pi$ and $\int_{-\pi}^{\pi} (1 - \cos u) K_m(u)\, du \sim \pi^2/2m^2$. To obtain the first result observe that $k_m \sim \pi^{-1} \sum_1^m w_m(j)^2$ and

$$\begin{aligned} w_m(j) = \pi^{-1} \{ &\sin[(j-1)\pi/(m+2)] \\ &+ \pi(m+2)^{-1}(m + 3 - j) \cos[(j-1)\pi/(m+2)] \} + O(m^{-1}) \end{aligned}$$

uniformly in $1 \leqslant j \leqslant m$, so that

$$\begin{aligned} \pi^3 k_m &\sim m \int_0^1 \{\sin(\pi u) + \pi(1-u)\cos(\pi u)\}^2\, du \\ &= m\pi^{-1} \int_0^\pi \{\sin u - u \cos u\}^2\, du = m(2\pi^2 + 15)/12. \end{aligned}$$

(iii) *Weierstrass singular integral.* Let $I = (0, \infty)$ and $r_0 = 0$, and define

$$K_r(u) = (2\pi)^{-1} \left[1 + 2 \sum_1^\infty e^{-rj^2} \cos ju \right] \geqslant 0.$$

In this case $k_r \sim (8\pi r)^{-1/2}$ and $\int_{-\pi}^{\pi} (1 - \cos u) K_m(u)\, du \sim r$.

4. Comparison of Kernel and Singular Integral Estimates

The estimators based on the Rogosinski and Korovkin singular integrals are considerably more easy to construct and update than a typical kernel estimator when the sample size is large. This ease of computation derives from the simple formula

$$\hat{f}_n(x; m) = (2\pi)^{-1} \left[1 + 2 \sum_{1}^{m} w_m(j)(\hat{a}_j \cos jx + \hat{b}_j \sin jx) \right],$$

where $\hat{a}_j = n^{-1} \sum_{i=1}^{n} \cos jX_i$ and $\hat{b}_j = n^{-1} \sum_{i=1}^{n} \sin jX_i$. If the value of m is chosen in an optimal way, the expansion (24) can be minimised, and we may write

$$\min_m \int_{-\pi+\varepsilon}^{\pi-\varepsilon} E[\hat{f}_n(x; m) - f(x)]^2 \, dx \sim C_0 C(f) \, n^{-4/5},$$

where the constant C_0 depends only on the singular integral type, and

$$C(f) = \left[\int_{-\pi+\varepsilon}^{\pi-\varepsilon} f(x) \, dx \right]^{4/5} \left[\int_{-\pi+\varepsilon}^{\pi-\varepsilon} f''(x)^2 \, dx \right]^{1/5}.$$

The same expansion may be achieved for kernel estimators if the window size is chosen optimally, and the values of C_0 are compared in the table. The kernel $K_0(u) = 3(5 - u^2)/20\sqrt{5}$ if $|u| \leqslant \sqrt{5}$; 0 if $|u| > \sqrt{5}$ is optimal in the sense of Epanechnikov [4]. See Table I.

The PMISE for the Rogosinski estimator is *smaller than that using the most efficient of the kernel estimators*. In interpreting this statement, however, it should be borne in mind that the PMISE formulae do not extend to the case $\varepsilon = 0$, and in some circumstances it may be preferable to obtain

TABLE I

Comparison of Optimal PMISE

Estimator type	C_0
Kernel; $K(u) = \frac{1}{2}$, if $\lvert u \rvert \leqslant 1$, 0 otherwise	0.4626
Kernel; $K(u) = (2\pi)^{-1/2} e^{-u^2/2}$	0.4542
Kernel; $K(u) = K_0(u)$	0.4364
Rogosinski singular integral	0.4123
Korovkin singular integral	0.4686
Weierstrass singular integral	0.4542

an estimate of the density on $(-\infty, \infty)$ by using a kernel estimate. For example, in the case of the Rogosinski or Korovkin estimators we have

$$\begin{aligned}\int_{-\pi}^{\pi} [E\hat{f}_n(x; m) - f(x)]^2\, dx &= \pi^{-1} \sum_{1}^{m} [1 - w_m(j)]^2 (a_j^2 + b_j^2) \\ &\quad + \pi^{-1} \sum_{m+1}^{\infty} (a_j^2 + b_j^2) \\ &\geqslant \pi^{-1} \sum_{m+1}^{\infty} (a_j^2 + b_j^2) \\ &= [f(\pi) - f(-\pi)]^2/m\pi + o(m^{-1}),\end{aligned}$$

where $a_j = \int_{-\pi}^{\pi} \cos ju\, f(u)\, du$ and $b_j = \int_{-\pi}^{\pi} \sin ju\, f(u)\, du$. Since

$$\int_{-\pi}^{\pi} \operatorname{var}[\hat{f}_n(x; m)]^2\, dx \sim \text{const} \times m/n,$$

the MISE will not converge at a faster rate than $O(n^{-1/2})$ unless $f(\pi) = f(-\pi)$.

Finally, we would like to stress that the Rogosinski and Korovkin estimators are guaranteed to be nonnegative.

Acknowledgment

The detailed suggestions of the referee have led to improvements in presentation. I am grateful for his comments.

References

[1] Anderson, G. L. and De Figueiredo, R. J. P. (1980). An adaptive orthogonal-series estimator for probability density functions. *Ann. Statist.* **8** 347–376.

[2] Bosq, D. and Bluez, J. (1978). Etude d'une classe d'estimateurs non-parametrique de la densité. *Ann. Inst. H. Poincaré Sect. A (N.S.)* **14** 479–498.

[3] Butzer, P. L. and Nessel, R. J. (1971). *Fourier Analysis and Approximation*, Vol. I. Academic Press, New York.

[4] Carslaw, H. S. (1930). *Introduction to the Theory of Fourier's Series and Integrals.* Dover, New York.

[5] Cencov, N. N. (1962). Evaluation of an unknown distribution density from observations. *Soviet Math. Dokl.* **3** 1559–1562.

[6] Elderton, W. P. (1953). *Frequency Curves and Correlation*, 4th ed. Cambridge Univ. Press, London.

[7] Epanechnikov, V. A. (1969). Nonparametric estimation of a multivariate probability density. *Theory Probab. Appl.* **14** 153–163.

[8] HALL, P. (1981). On trigonometric series estimates of densities. *Ann. Statist.*
[9] KRONMAL, R. AND TARTER, M. (1968). The estimation of probability densities and cumulatives by Fourier series methods. *J. Amer. Statist. Assoc.* **63** 925–952.
[10] LANCZOS, C. (1956). *Applied Analysis*. Prentice Hall, Englewood Cliffs, N.J.
[11] PETROV, V. V. (1976). *Sums of Independent Random Variables*. Springer-Verlag, Berlin.
[12] VAN RYZIN, J. (1966). Bayes risk consistency of classification procedures using density estimation. *Sankhyā Ser. A* **28** 261–270.
[13] SCHWARTZ, S. C. (1967). Estimation of probability density by orthogonal series. *Ann. Math. Statist.* **38** 1261–1265.
[14] TARTER, M. E. (1979). Trigonometric maximum likelihood estimation and application to the analysis of incomplete survival information. *J. Amer. Statist. Assoc.* **74** 132–139.
[15] TARTER, M. E. AND RAMAN, S. (1972). A systematic approach to graphical methods in biometry. In *Proc. Sixth Berkeley Symp. Math. Statist. and robability* **IV** 199–222.
[16] WAHBA, G. (1975). Optimal convergence properties of variable knot, kernel, and orthogonal series methods for density estimation. *Ann. Statist.* **3** 15–20.
[17] WAHBA, G. (1977). Optimal smoothing of density estimates. In *Classification and Clustering* (J. van Ryzin, Ed.), pp. 423–457. Academic Press, New York.
[18] WALTER, G. G. (1980). Properties of Hermite series estimation of probability density. *Ann. Statist.* **8** 454–455.
[19] WALTER, G. G. AND BLUM, J. R. (1979). Probability density estimation using delta sequences. *Ann. Statist.* **7** 328–340.
[20] WATSON, G. S. (1969). Density estimation by orthogonal series. *Ann. Math. Statist.* **40** 1496–1498.
[21] ZYGMUND, A. (1959). *Trigonometric Series*, Vol. I. Cambridge Univ. Press, Cambridge.